



TRANSLATION OF ANNEXES

retrieved (selected depending on a situation) from a huge number of chunks (mass of meanings: words, phrases, stereotyped expressions, collocations, sentence structures, and sentences) accumulated in an individual brain based on the relation to the contextual words before and after the word concerned and the background knowledge associated with the phrase itself.

[0008] Therefore, natural translations as sentences in the target language can be obtained by seeking corresponding translations of the original sentence after recognizing the contextual relations to the words before and after the word concerned, the background knowledge associated with the phrases themselves, and the units (that is to say, chunks) that people retrieve when they speak or write in their native language in the same way as when people speak or write in their native language. However, actual circumstances were that creation of translated sentences by seeking corresponding translations by chunks was not yet realized in the machine translation system because it is not easy to identify the context or the background knowledge associated with phrases themselves, and punctuation of chunks is obscure.

[0009] This invention was established taking the abovementioned facts into consideration, and has a purpose to obtain corresponding translation-identifying systems, corresponding translation identifying methods, and programs.

Measures to work out the assignment

[0010] In order to achieve the abovementioned purpose, the translation identifying system relating to the invention according to Claim 1 comprises a storage means for storing a plurality of natural sentences composed of a plurality of words in the source language by correlating the natural sentences with the translations in the target language, a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in a source language from a plurality of natural sentences in said source language stored in said storage means, search means for searching natural sentences extracted by retrieval of said retrieval means for said phrase to be translated and frequently appearing translations appearing in the same sentence of said source language, and a translation identifying means for searching translated sentences of natural sentences containing each of said phrase to be translated and specific frequently appearing phrases for frequently appearing translations of phrase to be translated by referring to said specific frequently appearing phrases found by said search means and existing in said original sentences and referring to translated sentences of natural sentences containing each of said phrase to be translated among natural sentences extracted by retrieval of said retrieval means and for

identifying found frequently appearing translations as translations of said phrases to be translated in said original sentences.

[0011] In the invention according to Claim 1, natural sentences composed of a plurality of words in the source language are stored in the storage means by correlating the natural sentences with the translations in the target language.

The natural sentence relating to this invention is at least one of a sentence, a clause, a phrase, a stereotyped expression, and a collocation that are not subjected to editing or processing, such as division by word unit and extraction of multivocal words as in a dictionary in a conventional machine translation system, eliminates the necessity of providing a multivocal word with many corresponding translations as in the case of registration in a dictionary by word unit by storing these natural sentences by correlating these sentences with the translations in the target language (when a multivocal word is correlated with many translations, there is a possibility of erroneous selection because a suitable translation must be chosen) and can transform the translation corresponding to a natural sentence to natural sentences in the target language.

Moreover, words and their corresponding translations in addition to the natural sentences may be stored in the storage means in this invention.

[0012] In the invention according to Claim 1, a natural sentence that contains phrases to be translated in the original sentence described in the source language (any one of a sentence, a clause, a phrase, a stereotyped expression and a collocation) is retrieved by the retrieval means from a plurality of natural sentences in the source language stored in the storage means.

The phrase to be translated that means the phrase especially to be translated in the original sentence may be a word or may be composed of a plurality of words.

For example, when the translation identifying system relating to this invention is used by a person (translator) as an electronic dictionary for translation work, the phrase to be translated is specified by the translator.

Moreover, for example, when the translation identifying system relating to this invention is used as a part of a machine translation system or an automatic interpretation system, the phrase to be translated is specified by the machine translation system or the automatic interpretation system that performs machine translation using the corresponding translation identified by the translation identifying system relating to this invention.

In addition, the original sentence containing the phrase to be translated may be also specified by a user (translator, etc.) or may be specified automatically (for example, a sentence or a clause containing the phrase to be translated is automatically identified as the original sentence).

Since the retrieval means retrieves a natural sentence containing a phrase to be translated, this retrieval results in extraction of a natural sentence containing a translation of the phrase to be translated in the corresponding translation.

[0013] In the invention according to Claim 1, a search means searches the natural sentences extracted by the retrieval of the retrieval means for the phrases to be translated and the frequently appearing phrases that frequently appear in the same sentences of the source language, a translation identifying means identifies translations of at least said phrase to be translated in translated sentences of natural sentences containing each of alternative phrases identified by said identifying means and said phrase to be translated among natural sentences extracted by retrieval of said retrieval means, as translations of at least said phrase to be translated in said original sentences.

[0014] When a phrase to be translated and specific frequently appearing phrases that frequently appear (called correlation) in the same sentence in the source language exist in the original sentence, there is a high probability that the translation of the phrase to be translated coincides with the translation of the phrase to be translated in the corresponding translation of a natural sentence containing each of the phrases to be translated and specific frequently appearing phrases. However, although there is a high probability that the phrases to be translated are identical to a suitable translation of the phrases to be translated in the corresponding translation of a natural sentence containing each of the specific frequently appearing phrases, there also is a possibility that natural sentences having unsuitable translations of phrases to be translated are mixed in the abovementioned natural sentences stored in the storage means.

Based on this, in the invention according to Claim 1, since frequently appearing phrases having a high correlation with phrases to be translated are searched for, frequently appearing translations for phrases to be translated and phrases to be translated in the translation of a natural sentence containing each of the specific frequently appearing phrases are found by referring to the specific frequently appearing phrases existing in the original sentence among the recognized frequently appearing phrases and the translation of a natural sentence containing each of the phrases to be translated. Therefore, suitable translations of phrases to be translated in the original sentence (corresponding translation where there is a high probability that natural translations are obtained as sentences in the target language from the original sentence) can be obtained based on the frequently appearing phrases even if it is difficult to identify (select) a suitable natural sentence based only on the degree of coincidence.

[0015] It is difficult to identify the suitable translation taking the contexts before and after a word into consideration because there are many options depending on the situation. In the

invention according to Claim 1, the corresponding translation of a phrase to be translated is identified by focusing on the phrase that frequently appears in the same text at the same time and by referring to the corresponding translation sentence of a natural sentence where this phrase and the phrase to be translated appear at the same time. Therefore, it is not necessary to take the contexts before and after the original sentence into consideration. As a result, a suitable translation of a multivocal word can be obtained roughly considering the contexts before and after the original sentence.

It is possible to search for the abovementioned frequently appearing phrases by registering phrases that frequently appear in the same text of the source language in a table and referring to the table. In the invention according to Claim 1, since the frequently appearing phrases are found based on the natural sentences extracted by the retrieval of the retrieval means, there are two advantages wherein time and labor for preparing the abovementioned table can be saved, and the storage capacity required for storing the table can also be saved.

[0016] The translation identifying system relating to the invention according to Claim 2 comprises a storage means for storing a plurality of natural sentences, of original sentences composed of a plurality of words, correlated with translated sentences in a target language, a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in a source language from a plurality of natural sentences in said source language stored in said storage means, an identifying means for identifying alternative phrases existing in said original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said retrieval means, an identifying means for identifying alternative phrases existing in said original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said retrieval means, a translation identifying means for identifying translations of at least said phrase to be translated in translated sentences of natural sentences containing each of alternative phrases identified by said identifying means and said phrase to be translated among natural sentences extracted by retrieval of said retrieval means, as translations of at least said phrase to be translated in said original sentences.

[0017] In common with the invention as defined in claim 1, multiple natural sentences in the source language, respectively composed of multiple words, are stored in memory in the invention as defined in claim 2, corresponding to the parallel version in a target language out of which natural sentences containing the targeted phrases in the original text in a source language are

retrieved by the search tool.

A natural sentence where a specific phrase in the original sentence is replaced with another phrase is sometimes extracted by the retrieval of the retrieval means. In this case, if meanings of the original sentence and the natural sentence are similar, the specific phrase is considered to have a replaceable relation to the other phrase.

Even if a plurality of natural sentences is extracted by the retrieval of the retrieval means, when a natural sentence where a specific phrase for the original sentence is replaced with the other phrase (alternative phrase) having a replaceable relation to the specific phrase exist in these natural sentences, it is preferable to select this natural sentence because there is a high probability that this natural sentence and the original sentence have the same meaning.

[0018] Based on this, in the invention according to Claim 2, an alternative phrase that is replaceable with a targeted phrase extracted by the retrieval of the retrieval means and not contained in the natural sentence existing in the original sentence is identified by the identifying means, and the translation identifying means identifies the translation of the alternative phrase identified by the identifying means and of at least the phrase to be translated in the corresponding translation sentence of a natural sentence containing each of the phrases to be translated among the natural sentences extracted by the retrieval of the retrieval means as the translation of at least the phrase to be translated in the original sentence.

Thus, based on an alternative phrase, suitable translations of phrases to be translated in original sentences having a high probability of obtaining natural translated sentences as sentences in a target language from the original sentences in a source language can be obtained.

[0019] Identification of an alternative phrase by the identifying means relating to the invention according to Claim 2 may be performed by registering phrases having replaceable relations with each other in a table and referring to this table. For example the identification may be performed by retrieving a natural sentence containing a targeted phrase from a plurality of natural sentences stored in the storage means, retrieving a natural sentence having a sentence structure identical to the natural sentence extracted by this retrieval from the natural sentences stored in the storage means, and identifying the phrase replaced with the targeted phrase in the natural sentence extracted by this retrieval as the alternative phrase as described in Claim 3.

Also, in this case, time and labor for preparing the abovementioned table can be saved, and the storage capacity required for storing the table can be also saved.

[0020] Since an immeasurable amount of background knowledge associated with phrases themselves having various properties depending on these phrases exist, it is difficult to extract all of them to classify into categories.

In the inventions according to Claims 2 and 3, since a phrase replaceable in the same sentence structure is identified as a phrase in the same category, it is possible to obtain suitable translations roughly based on the background knowledge associated with phrases themselves without obtaining background knowledge of entire phrases to classify them into categories.

[0021] In the invention according to Claim 1 or 2, as described in, for example, Claim 4, it is preferable that a translation identifying means evaluates a degree of coincidence between natural sentences extracted by retrieval of said retrieval means and said original sentence and identifies translations of at least said phrase to be translated in translated sentences of natural sentences selected on the basis of said evaluated degree of coincidence, as translations of at least said phrase to be translated in said original sentence.

Thus a natural sentence similar to the original sentence among natural sentences extracted by the retrieval of the retrieval means (for example, a natural sentence used with the same meaning as the multivocal word existing in the original sentence) is selected with high probability as a natural sentence with a high degree of coincidence, a natural translation corresponding to the selected natural sentence is obtained, and the translation of, at least, the phrase to be translated in the translations are identified as the translation of the phrase to be translated in the original sentence (it is a matter of course that the translation of a phrase other than the phrase to be translated in the corresponding translation can be identified as the translation of the phrase in the original sentence depending on the degree of coincidence with the original sentence).

[0022] Thus the invention according to Claim 4, considering the great technical difficulty of identifying punctuation of chunks in the original sentence, stores a plurality of natural sentences in the source language that correspond to the translations in the target language, selects a natural sentence having a high degree of coincidence with the original sentence in the source language from the stored natural sentences, and identifies the translation in the natural translations corresponding to the selected natural sentences as the translation of, at least, the phrase to be translated in the original sentence. Therefore the invention enables the obtaining of natural translations equal to the translation corresponding to the original sentence performed generally by chunk as a result that it is not necessary to identify the punctuation of the chunk in the original

sentence and enables the obtaining of suitable translations of the phrases to be translated in the original sentence (the translation having a high probability of obtaining from the original sentence a natural translation as a sentence in the target language).

Moreover, since the invention according to Claim 4 determines a degree of coincidence between a natural sentence extracted by the retrieval of the retrieval means and the original sentence, and can obtain the suitable translation by selecting a natural sentence based on the determined degree of coincidence, it is not necessary to perform complicated processing, such as identification of a part of speech and an analysis of sentence structure as in the conventional machine translation system, and can simplify the processing.

[0023] The parallel translation identifying system related to the invention as defined in claim 5 consists of the following tools: the memorizing means storing multiple natural sentences composed of multiple words in a source language coordinating with the parallel sentences in a target language; the retrieval tool searching natural sentences containing the object phrases to be rendered in the original text in the source language out of the multiple natural sentences in the source language retained in the above memory; and the parallel translation identifying system processing the minimal translation target phrase in the natural parallel version selected based on the total value of the first evaluated value computed as the minimal object phrase in the original text by respectively adding the first evaluated value corresponding to the small distance between the phrase to be rendered and the respective words in the original text, identifying the words matching the original from among the respective words in the natural sentence in question about the natural one extracted through the retrieval tool, and by computing by the natural sentence for which the total value of the said first evaluated value added to the corresponding words were recognized.

[0024] The invention as defined in claim 5 is equipped with the same memory and search means as defined in claims 1 and 2 with the first evaluated value individually assigned corresponding to the small distance between the respective words and the phrase to be rendered in the original text, recognizing the matching words with the original from among the respective words in the relevant natural sentences extracted by the retrieval unit, computing by the natural sentence the total value of the first value given to the matching words recognized, based on which identifying the parallel version of the minimal object phrases in the selected natural sentence as the corresponding translation of the minimal target phrases in the original text. Therefore, on the basis of the total

value of the first evaluated value, being capable of selecting the natural sentence containing more words located in the proximity of the target phrases to be rendered in the original text, i.e., a natural sentence with the high possibility of containing a collocation of phrases composed of the target phrases and words in the vicinity to be rendered in the original text (a natural sentence assumed with higher degree of similarity in the alignment sequence of the words compared with the original) can be chosen, thus, acquiring an appropriate parallel translation of the phrases in the original to be rendered with the high possibility of producing a natural rendition as a sentence in the target language from the original text in the source language (parallel rendition with higher possibility of natural translation from the original text into the target language).

[0025] In addition, in the invention as defined in claim 5, it will be preferable to configure the parallel translation identifying means in such a way as defined in claim 6, if there are multiple sentences from the computed result of the first evaluated value, about the respective natural sentences of such multiplicity, that the parallel translation of the minimal phrases to be rendered in the parallel translation of the natural sentence selected based on the computed total value of the second evaluated value can be identified as the parallel translation for the minimal phrases in the original text, respectively assigning the second evaluated value according to the small distance in the above natural sentence from the object phrases for rendering in the natural sentence for the matching words in the said sentence, computing the total value of the second evaluated value given to the corresponding words aforementioned by the natural sentence.

Thus, if there are multiple sentences from the computed result of the first evaluated value, on the basis of the total value of the second evaluated value, the natural sentence containing a larger number of the words coinciding with the original sentence and containing the mismatched words existing between the coinciding words as small as possible, i.e., a natural sentence with the high possibility of containing a collocation of phrases composed of the target phrases and words in the vicinity to be rendered in the original text (a natural sentence assumed with higher degree of similarity in the alignment sequence of the words compared with the original) can be chosen, thus, acquiring an appropriate parallel translation of the phrases in the original to be rendered with the high possibility of producing a natural rendition as a sentence in the target language from the original text in the source language (parallel rendition with higher possibility of natural translation from the original text into the target language).

[0026] The parallel translation identifying system related to the invention as defined in claim 7

consists of the following means and unit: the memorizing means storing multiple sentences in the source language composed of multiple words coordinating the rendered parallel sentences of the target language; the search unit retrieving the natural sentence containing the target phrases to be rendered in the original text of the source language from the multiple natural sentences in the source language stored in the above memory; and the parallel translation identifying system recognizing the rendition of the said minimal parallel phrases in the parallel natural sentences selected based on the total value of the third evaluated value computed as the parallel translation of the minimal parallel phrases in the original text in question, searching, to the natural sentence extracted by the search unit, from among the anterior word groups aligned toward the front side from the object phrases to be rendered in the original text, whether the specific anterior words non-retrieved and with minimum distance from the object phrases to be rendered are located toward the front of the object phrases to be rendered in the natural sentence or the anterior words extracted in the preceding search within the prescribed number of words, repeatedly assigning the third evaluated value according to the minimal distance between the specific anterior words upon detection and the object phrases to be rendered or the anterior words extracted in the preceding search until the non-retrieved anterior words have ceased to exist in the aforementioned original text, and concomitantly searching, to the natural sentence extracted by the search unit, from among the posterior word groups aligned toward the back from the object phrases to be rendered in the original text, whether the specific posterior words non-retrieved and with minimum distance from the object phrases to be rendered are located toward the back of the object phrases to be rendered in the natural sentence or the posterior words extracted in the preceding search within the prescribed number of words, repeatedly assigning the third evaluated value according to the minimal distance between the specific posterior words upon detection and the object phrases to be rendered or the posterior words extracted in the preceding search until the non-retrieved posterior words have ceased to exist in the aforementioned original text, thus, computing the total value of the third evaluated value by the natural sentence.

[0027] The invention as defined in claim 7 is equipped with the same memorizing means and the search unit as defined in claims 1 and 2, with the parallel translation identifying system, searching, to the natural sentence extracted by the search unit, from among the anterior word groups aligned toward the front side from the object phrases to be rendered in the original text, whether the specific anterior words non-retrieved and with minimum distance from the object phrases to be

rendered are located toward the front of the object phrases to be rendered in the natural sentence or the anterior words extracted in the preceding search within the prescribed number of words, repeatedly assigning the third evaluated value according to the minimal distance between the specific anterior words upon detection and the object phrases to be rendered or the anterior words extracted in the preceding search until the non-retrieved anterior words have ceased to exist in the aforementioned original text, and concomitantly searching, to the natural sentence extracted by the search unit, from among the posterior word groups aligned toward the back from the object phrases to be rendered in the original text, whether the specific posterior words non-retrieved and with minimum distance from the object phrases to be rendered are located toward the back of the object phrases to be rendered in the natural sentence or the posterior words extracted in the preceding search within the prescribed number of words, repeatedly assigning the third evaluated value according to the minimal distance between the specific posterior words upon detection and the object phrases to be rendered or the posterior words extracted in the preceding search until the non-retrieved posterior words have ceased to exist in the aforementioned original text, thus, computing the total value of the third evaluated value by the natural sentence. Since the object phrases to be rendered in the object sentence of the natural sentence are at least identified based on the total value of the third evaluated value as the object phrases to be rendered in the original text at least, the similarity in the word alignment sequence is quite high compared to the original text on the basis of the total value of the third evaluated value, and concomitantly, a natural sentence with the least number of discordant words, i.e., the natural sentence with the original text excluding discordant words as much as possible (status of tighter group of words) can thus be chosen and created with a higher possibility of obtaining a natural parallel version with more appropriate phrases as a sentence of the target language rendered from the original text in the source language.

[0028] The parallel translation identifying system related to the invention as defined in claim 8 consists of the following means and unit: the memorizing means storing multiple sentences in the source language composed of multiple words coordinating the rendered parallel sentences of the target language; the search unit retrieving the natural sentence containing the target phrases to be rendered in the original text of the source language from the multiple natural sentences in the source language stored in the above memory; and the parallel translation identifying system recognizing at least the parallel version of the object phrases to be rendered in the parallel natural

sentence selected based on the total value computed from the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side as at least the parallel version for the object phrases to be rendered in the original text, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the first anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the front side corresponding to the distance from the object phrases to be rendered when the specific anterior words are detected or the from anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the first anterior group of words, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the first posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the back side corresponding to the distance from the object phrases to be rendered when the specific posterior words are detected or the from posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the first posterior group of words.

[0029] The invention as defined in claim 8 is equipped with the same memorizing means and the search unit as defined in claims 1 and 2, with the parallel translation identifying system recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the total value computed from the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side as at least the parallel version for the object phrases to be rendered in the original text, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the

natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the first anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the front side corresponding to the distance from the object phrases to be rendered when the specific anterior words are detected or the from anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the first anterior group of words, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the first posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the back side corresponding to the distance from the object phrases to be rendered when the specific posterior words are detected or the from posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the first posterior group of words. Thus, the similarity in the word alignment sequence is quite high compared to the original text on the basis of the total value computed from the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side, and concomitantly, a natural sentence with the least number of discordant words, i.e., the natural sentence with the original text excluding discordant words as much as possible (status of tighter group of words) can be chosen and created with a higher possibility of obtaining a natural parallel version with more appropriate phrases as a sentence of the target language rendered from the original text in the source language.

[0030] Moreover, the parallel translation identifying system in the invention as defined in claim 8 as in 9, for example, is located on the front compared to the object phrases to be rendered in the original text toward the natural sentence extracted by the search unit, exploring whether the specific front words non-retrieved and with minimal distance from the object phrases to be rendered among the second anterior group of words excluding the anterior words adjacent to the object phrases to be rendered are aligned on the front of the object phrases to be rendered in the natural sentence or anterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up 1 to the distance between the specific front

words when discovered and the object phrases to be rendered or the front words extracted in the preceding search or adding the number of words by the second interval in relation to the corresponding front side to the distance added to the interval computed regarding the front-side words discovered in the preceding search until the non-retrieved front words have ceased to exist in the original text, concomitantly selecting as the final recognition the result of a shorter distance between the respective words from among the number of words by the first and second intervals regarding the front side respectively, and also, toward the natural sentence extracted by the search unit, exploring whether the specific back-side words non-retrieved and with minimal distance from the object phrases to be rendered among the second posterior group of words excluding the posterior words adjacent to the object phrases to be rendered are aligned on the back of the object phrases to be rendered in the natural sentence or posterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up 1 to the distance between the specific back-side words when discovered and the object phrases to be rendered or the back-side words extracted in the preceding search or adding the number of words by the second interval in relation to the corresponding back side to the distance added to the interval computed regarding the back-side words discovered in the preceding search until the non-retrieved back-side words have ceased to exist in the original text, concomitantly selecting as the final recognition the result of a shorter distance between the respective words from among the number of words by the first and second intervals regarding the back side respectively, performing a comprehensive evaluation by adding up the final recognition on the anterior and posterior sides, thus, preferably configuring and identifying at least the parallel version of the object phrases to be rendered in the parallel sentence of the natural sentence of the result with shorter intra-word distance in the computed integrated recognition as the parallel translation at least of the target phrases to be rendered in the original text.

As shown in the above, by using the number of words by the second distance in combination with those by the first distance, the degree of similarity in the alignment sequence of the words compared to the original text can be justly evaluated by the number of words by the second distance also regarding the natural sentence the sequence of the part of which the words have been changed toward the original text, thus, improving the probability of obtaining more a appropriate parallel version of the object phrases to be rendered in the original text.

[0031] The translation identifying method relating to the invention according to Claim 10

comprises a first step for retrieving natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language, a second step for searching the natural sentences extracted by the retrieval of the retrieval means in the aforementioned first step for the aforementioned phrases to be translated and the frequently appearing phrases that frequently appear in the same sentences of the source language, and a third step for searching for the aforementioned phrases to be translated and the frequently appearing translations for the phrases to be translated in the corresponding translation sentences of the natural sentences containing each of the aforementioned specific frequently appearing phrases and identifying the found frequently appearing translations as the translations of the aforementioned phrases to be translated in the original sentence by referring to the specific frequently appearing phrases found in the aforementioned second step and existing in the aforementioned original sentence among the natural sentences extracted by the retrieval in the aforementioned first step and referring to the corresponding translation sentences of the aforementioned natural sentences each of which contains the aforementioned phrases to be translated. Therefore in the same manner as the invention according to Claim 1, the suitable translations of the phrases to be translated in the original sentence (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language) can be obtained.

[0032] The translation identifying method relating to the invention according to Claim 11 comprises a first step for retrieving natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language, a second step for identifying alternative phrases existing in said original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said first step, and a third step for identifying translations of at least said phrase to be translated in translated sentences of natural sentences containing each of alternative phrases identified by said second step and said phrase to be translated among natural sentences extracted by retrieval of said first step, as translations of at least said phrase to be translated in said original sentences. Thus, in the same manner as the invention according to Claim 2, suitable translations of phrases to be translated in the original sentence (the translation having high

probability of obtaining from the original sentence a natural translation as a sentence in the target language) can be obtained.

[0033] The parallel translation identifying system related to the invention as defined in claim 12 consists of two steps: the first step to search the natural sentence composed of object phrases to be rendered in the original text in a source language from out of the natural sentences in a source language composed of multiple words respectively stored in multiples in memory in response to the parallel version of the target language; and the second step to identify the parallel version of the object phrases to be rendered at least in the parallel sentence of the natural sentence selected based on the total value of the first evaluated value computed at least as the parallel version of the object phrases to be rendered in the original text, respectively granting the first evaluated value corresponding to the small distance in the original text from the object phrases to be rendered in the original text toward the respective words in the original text, recognizing the words concordant with the original text from among the respective words in the relevant natural sentence in relation to the natural sentence extracted by the search process in the first step, thus, computing by the natural sentence the total value of the first evaluated value granted to the corresponding words identified. Therefore, in the same way as the invention as defined in claim 5, an appropriate parallel translation can be obtained for the object phrases to be rendered in the original text (i.e., parallel version with a higher possibility of acquiring a natural sentence as a sentence in the target language from the original text).

[0034] The parallel translation identifying system related to the invention as defined in claim 13 consists of two steps: the first step to search the natural sentence composed of object phrases to be rendered in the original text in a source language from out of the natural sentences in a source language composed of multiple words respectively stored in multiples in memory in response to the parallel version of the target language; and the second step to recognize the rendition of the said minimal parallel phrases in the parallel natural sentences selected based on the total value of the third evaluated value computed as the parallel translation of the minimal parallel phrases in the original text in question, searching, to the natural sentence extracted by the search process in the first step, from among the anterior word groups aligned toward the front side from the object phrases to be rendered in the original text, whether the specific anterior words non-retrieved and with minimum distance from the object phrases to be rendered are located toward the front of the object phrases to be rendered in the natural sentence or the anterior words extracted in the

preceding search within the prescribed number of words, repeatedly assigning the third evaluated value according to the minimal distance between the specific anterior words upon detection and the object phrases to be rendered or the anterior words extracted in the preceding search until the non-retrieved anterior words have ceased to exist in the aforementioned original text, and concomitantly searching, to the natural sentence extracted by the search process in the first step, from among the posterior word groups aligned toward the back from the object phrases to be rendered in the original text, whether the specific posterior words non-retrieved and with minimum distance from the object phrases to be rendered are located toward the back of the object phrases to be rendered in the natural sentence or the posterior words extracted in the preceding search within the prescribed number of words, repeatedly assigning the third evaluated value according to the minimal distance between the specific posterior words upon detection and the object phrases to be rendered or the posterior words extracted in the preceding search until the non-retrieved posterior words have ceased to exist in the aforementioned original text, thus, computing the total value of the third evaluated value by the natural sentence. Therefore, in the same way as the invention as defined in claim 7, an appropriate parallel translation can be obtained for the object phrases to be rendered in the original text (i.e., parallel version with a higher possibility of acquiring a natural sentence as a sentence in the target language from the original text).

[0035] The parallel translation identifying system related to the invention as defined in claim 14 consists of two steps: the first step to search the natural sentence composed of object phrases to be rendered in the original text in a source language from out of the natural sentences in a source language composed of multiple words respectively stored in multiples in memory in response to the parallel version of the target language; and the second step to recognize at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the total value computed from the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side as at least the parallel version for the object phrases in the original text, searching, regarding the natural sentence extracted by the search process in the first step, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the

first anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the front side concordant with the interval from the object phrases to be rendered upon detection of the specific anterior words and from the anterior words extracted in the preceding search process till the non-retrieved anterior words have ceased to exist in the first anterior group of words, and concomitantly, searching, regarding the natural sentence extracted through the process by the search process in the first step, to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the first posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the back side concordant with the interval from the object phrases to be rendered upon detection of the specific posterior words and from the posterior words extracted in the preceding search process till the non-retrieved posterior words have ceased to exist in the first posterior group of words. Therefore, in the same way as the invention as defined in claim 8, an appropriate parallel translation can be obtained for the object phrases to be rendered in the original text (i.e., parallel version with a higher possibility of acquiring a natural sentence as a sentence in the target language from the original text).

[0036] The program relating to the invention according to Claim 15 allows a computer connected to a storage means for storing a plurality of natural sentences in the source language composed of a plurality of words having corresponding translations in the target language to function as a retrieval means for retrieving the natural sentences containing the phrases to be translated in the original sentence in the source language from a plurality of natural sentences in the source language stored in the aforementioned storage means, a search means for searching the natural sentences extracted by the retrieval by the aforementioned retrieval means for the aforementioned phrases to be translated and the frequently appearing phrases that frequently appear in the same sentences of the source language, and a second translation identifying means for searching for the aforementioned phrases to be translated and the frequently appearing translations for the phrases to be translated in the corresponding translation sentences of the natural sentences each of which contains the aforementioned specific frequently appearing phrases

and for identifying the found frequently appearing translations as the translations of the aforementioned phrases to be translated in the aforementioned original sentence by referring to the specific frequently appearing phrases found by the aforementioned search means and existing in the aforementioned original sentence among the natural sentences extracted by the retrieval by the aforementioned retrieval means and referring to the corresponding translation sentences of the natural sentences each of which contains the aforementioned phrases to be translated.

The program relating to the invention according to Claim 15 is a program for allowing a computer connected to a storage means for storing a plurality of natural sentences in the source language composed of a plurality of words having corresponding translations in the target language (either a computer in which a storage means is built or a computer connected to another computer connected to a storage means through telecommunication line) to function as the abovementioned retrieval means, search means, and translation identifying means. Therefore, when the abovementioned computer runs the program relating to the invention according to Claim 15, the abovementioned computer functions as the translation identifying means according to Claim 1. Therefore in the same manner as the invention according to Claim 1, the suitable translations of the phrases to be translated in the original sentence (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language) can be obtained.

[0036] The program relating to the invention according to Claim 16 allows a computer connected to a storage means for storing a plurality of natural sentences in the source language composed of a plurality of words having corresponding translations in the target language to function as a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in a source language from a plurality of natural sentences in said source language stored in said storage means, an identifying means for identifying alternative phrases existing in said original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said retrieval means, and a translation identifying means for identifying translations of at least said phrase to be translated in translated sentences of natural sentences containing each of alternative phrases identified by said identifying means and said phrase to be translated among natural sentences extracted by retrieval of said retrieval means, as translations of at least said phrase to be translated in said original sentences.

The program relating to the invention according to Claim 16 is a program for allowing a computer

connected to the abovementioned storage means to function as the abovementioned retrieval means, search means, and translation identifying means. Therefore, when the abovementioned computer runs the program relating to the invention according to Claim 16, the computer functions as the translation identifying means according to Claim 2. Therefore, in the same manner as the invention according to Claim 2, the suitable translations of the phrases to be translated in the original sentence (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language) can be obtained.

[0037] The program relating to the invention according to Claim 17 allows a computer connected to a storage means for storing a plurality of natural sentences in the source language composed of a plurality of words having corresponding translations in the target language to function as a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in a source language from a plurality of natural sentences in said source language stored in said storage means, and a parallel translation identifying system for processing the minimal translation target phrase in the natural parallel version selected based on the total value of the first evaluated value computed as the minimal object phrase in the original text by respectively adding the first evaluated value corresponding to the small distance between the phrase to be rendered and the respective words in the original text, identifying the words matching the original from among the respective words in the natural sentence in question about the natural one extracted through the retrieval tool, and by computing by the natural sentence for which the total value of the said first evaluated value added to the corresponding words were recognized.

The program relating to the invention according to Claim 17 is a program for allowing a computer connected to the abovementioned storage means to function as the abovementioned retrieval means and translation identifying means. Therefore, when the abovementioned computer runs the program relating to the invention according to Claim 17, the computer functions as the translation identifying means according to Claim 5. Therefore, in the same manner as the invention according to Claim 5, the suitable translations of the phrases to be translated in the original sentence (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language) can be obtained.

[0038] The program relating to the invention according to Claim 18 allows the parallel translation identifying methodology to function recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the computed total value of the

third evaluated value as at least the parallel version for the object phrases to be rendered in the original text, making the methodology work together with other functions: the memorizing means storing multiple natural sentences composed of multiple words in source language by matching the parallel sentences in a target language and the search unit, by means of the connected computer, retrieving a natural sentence containing the object phrases to be rendered in the original text in a source language from among multiple natural sentences in a source language stored in the memory, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of granting the third evaluated value concordant with the small interval between the detected specific anterior words if detected and the object phrases to be rendered or the anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the original text, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of granting the third evaluated value concordant with the small interval between the detected specific posterior words if detected and the object phrases to be rendered or the posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the original text.

The program relating to the invention according to Claim 18 is a program for allowing a computer connected to the abovementioned storage means to function as the abovementioned retrieval means and translation identifying means. Therefore, when the abovementioned computer runs the program relating to the invention according to Claim 18, the computer functions as the translation identifying means according to Claim 7. Therefore, in the same manner as the invention according to Claim 7, the suitable translations of the phrases to be translated in the original sentence (the translation having high probability of obtaining from the original sentence a natural translation as

a sentence in the target language) can be obtained.

[0039] The program relating to the invention according to Claim 19 allows the parallel translation identifying methodology to function recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side as at least the parallel version for the object phrases to be rendered in the original text, making the methodology work together with other functions: the memorizing means storing multiple natural sentences composed of multiple words in a source language by matching the parallel sentences in a target language and the search unit, by means of the connected computer, retrieving a natural sentence containing the object phrases to be rendered in the original text in a source language from among multiple natural sentences in a source language stored in the memory, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the first anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the front side corresponding to the distance from the object phrases to be rendered when the specific anterior words are detected or the from anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the first anterior group of words, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the first posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the back side corresponding to the distance from the object phrases to be rendered when the specific posterior words are detected or the from posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the first posterior group of words.

The program relating to the invention according to Claim 19 is a program for allowing a computer connected to the abovementioned storage means to function as the abovementioned retrieval means and translation identifying means. Therefore, when the abovementioned computer runs the program relating to the invention according to Claim 19, the computer functions as the translation identifying means according to Claim 8. Therefore, in the same manner as the invention according to Claim 8, the suitable translations of the phrases to be translated in the original sentence (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language) can be obtained.

Effect of invention

[0040] As described above, this invention retrieves natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language, searches the natural sentences extracted by the retrieval for the phrases to be translated and the frequently appearing phrases that frequently appear in the same sentences of the source language, searches translated sentences of natural sentences containing each of the phrase to be translated and specific frequently appearing phrases for frequently appearing translations of phrase to be translated by referring to the specific frequently appearing phrases found by the search means and existing in the original sentences and referring to translated sentences of natural sentences containing each of the phrase to be translated and identifies found frequently appearing translations as translations of the phrases to be translated in the original sentences. Therefore, this invention has an excellent effect wherein the suitable translations of the phrases to be translated in the original sentence can be obtained.

[0041] Moreover this invention retrieves natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language, identifies alternative phrases existing in original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said retrieval means, and identifies translations of at least a phrase to be translated in translated sentences of natural sentences containing each of alternative phrase identified by said identifying means and phrase to be translated among natural sentences extracted by retrieval of said retrieval means, as translations of at least a phrase to be translated in original sentences.

Therefore, this invention has an excellent effect wherein the suitable translations of the phrases to be translated in the original sentence can be obtained.

[0042] Furthermore, this invention has an excellent advantageous effect capable of obtaining the most appropriate parallel version of the object phrases to be rendered in the original text since the parallel translation of the object phrases to be rendered at least in the parallel version of the natural sentence selected based on the total value of the first evaluated value computed is identified as at least the parallel version of the object phrases to be rendered in the original text, searching the natural sentence containing the object phrases to be rendered in the original text in a source language from out of the natural sentences in a source language respectively composed of multiple words stored in multiples in memory in response to the parallel sentence in a target language, respectively granting the first evaluated value according to the small interval in the original text between the object phrases to be rendered in the original text for the respective words in the original text, identifying the concordant words with the original text from among the respective words of the natural sentence in question regarding the natural sentence extracted by the search process, thus, computing by the natural sentence the total value of the first evaluated value granted to the corresponding words identified.

[0043] Additionally, this invention has an excellent advantageous effect capable of obtaining the most appropriate parallel version of the object phrases to be rendered in the original text since the parallel translation of the object phrases to be rendered at least in the parallel version of the natural sentence selected based on the total value of the third evaluated value computed is identified as at least the parallel version of the object phrases to be rendered in the original text, searching the natural sentence containing the object phrases to be rendered in the original text in a source language from out of the natural sentences in a source language respectively composed of multiple words stored in multiples in memory in response to the parallel sentence in a target language, searching, to the natural sentence extracted by the search process, from among the anterior word groups aligned toward the front side from the object phrases to be rendered in the original text, whether the specific anterior words non-retrieved and with minimum distance from the object phrases to be rendered are located toward the front of the object phrases to be rendered in the natural sentence or the anterior words extracted in the preceding search within the prescribed number of words, repeatedly assigning the third evaluated value according to the minimal distance between the specific anterior words upon detection and the object phrases to be rendered or the

anterior words extracted in the preceding search until the non-retrieved anterior words have ceased to exist in the aforementioned original text, and concomitantly searching, to the natural sentence extracted by the search process, from among the posterior word groups aligned toward the back from the object phrases to be rendered in the original text, whether the specific posterior words non-retrieved and with minimum distance from the object phrases to be rendered are located toward the back of the object phrases to be rendered in the natural sentence or the posterior words extracted in the preceding search within the prescribed number of words, repeatedly assigning the third evaluated value according to the minimal distance between the specific posterior words upon detection and the object phrases to be rendered or the posterior words extracted in the preceding search until the non-retrieved posterior words have ceased to exist in the aforementioned original text, thus, computing the total value of the third evaluated value by the natural sentence.

[0044] Additionally, this invention has an excellent advantageous effect capable of obtaining the most appropriate parallel version of the object phrases to be rendered in the original text since the parallel translation of the object phrases to be rendered at least in the parallel version of the natural sentence selected based on the integrated recognition computed from the number of words by the first distance regarding the front side and the number of words by the first distance regarding the back side is identified as at least the parallel version of the object phrases to be rendered in the original text, searching the natural sentence containing the object phrases to be rendered in the original text in a source language from out of the natural sentences in a source language respectively composed of multiple words stored in multiples in memory in response to the parallel sentence in a target language, exploring whether the specific front words non-retrieved and with minimal distance from the object phrases to be rendered among the first anterior group of words existing to the front to the object phrases to be rendered in the original text for the natural sentence extracted in the search in question are located on the front of the object phrases to be rendered in the natural sentence or anterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up the number of words by the first distance regarding the front side corresponding to the interval from the object phrases to be rendered or from the front words extracted in the preceding search when specific anterior words have been detected until the non-retrieved front words have ceased to exist in the first anterior group of words, and concomitantly, exploring whether the specific back words non-retrieved and with

minimal distance from the object phrases to be rendered among the first posterior group of words existing to the back to the object phrases to be rendered in the original text for the natural sentence extracted in the search in question are located on the back of the object phrases to be rendered in the natural sentence or posterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up the number of words by the first distance regarding the back side corresponding to the interval from the object phrases to be rendered or from the back words extracted in the preceding search when specific posterior words have been detected until the non-retrieved back words have ceased to exist in the first posterior group of words.

Brief description of the drawings

[0045]Figure 1 is a block schematic diagram showing a configuration of a personal computer related to the preferred embodiment.

Figure 2 is a flow chart showing the contents of the translation identifying processing.

Figure 3 is a table showing an example of natural sentences and the corresponding translations including a word “operation” registered in a translation database.

Figure 4 is a block diagram for describing a mode where this invention is applied to a computer system to which a client personal computer and a server are connected through a network.

Figure 5 is a block schematic diagram showing a configuration of a machine translation system to which this invention is applied.

The Best Mode of Embodiment of the Invention

[0046] Now one of the preferred embodiments of this invention is described in detail and referenced by drawings.

First, a mode where this invention is applied to an electronic dictionary system that provides a user with a translation of a specified phrase to be translated when the user specifies a phrase to be translated is described.

[0047] Figure 1 shows a personal computer (PC) 10 capable of functioning as the abovementioned electronic dictionary.

PC 10 comprises a CPU 10A, a ROM 10B, a RAM 10C and an I/O port 10D, which are connected with each other through a bus 10E composed of a data bus, an address bus, a control bus, etc., to configure the PC.

A display 12 composed of a CRT, an LCD, etc., a keyboard 14 through which a user enters data, etc., a mouse 16, a hard disk drive (HDD) 18, a CD-ROM drive 20 that reads data from CD-ROM 24, and a scanner 22 capable of reading paper data are connected to the I/O port 10D as various I/O devices.

[0048] A translation identifying program (corresponding to programs described in Claims 15 and 16) for allowing PC 10 to function as an electronic dictionary system is installed in HDD 18 of PC 10. A corresponding translation database (corresponding translation database) storing data that the aforementioned translation-identifying program uses for identifying a suitable translation is stored in this HDD 18.

There are several methods for installing (transferring) the translation-identifying program to a PC 10. For example, the translation-identifying program is installed by recording the translation-identifying program along with a setup program in CD-ROM 24 in advance, inserting the CD-ROM 24 into CD-ROM drive 20, and commanding CPU 10A to run the aforementioned setup program to read the translation-identifying program from CD-ROM 24 in sequence and to write the read translation-identifying program in HDD 18 in sequence.

[0049] The corresponding translation database can be stored in HDD 18 by, for example, recording the corresponding translation database in CD-ROM 24 in advance, and configuring the setup program so that the corresponding translation database is written in HDD 18 simultaneously when installing the translation-identifying program.

A large amount of text data of natural sentences composed of a plurality of words and described in

the source language (either one of sentences, clauses, phrases, stereotyped expressions and

an operation such as selecting the required icon displayed in, for example, the toolbar instead of using the context menu as described above.

When the abovementioned operation for specifying the output of the translation of the phrase to be translated is performed, CPU 10A of PC 10 runs the translation-identifying program to perform the translation identifying processing shown in Figure 2.

This translation identifying processing is processing to which the translation identifying method described in Claims 10 and 11 is applied. By performing this processing, PC 10 functions as an electronic dictionary system (the translation identifying system described in Claims 1 and 4).

[0053] In translation identifying processing, text data of a single original sentence (original sentence to be processed) containing the specified phrase to be translated are first acquired in Step 100, and then information for identifying the phrase to be translated in the acquired original sentence to be processed is acquired.

The original sentence to be processed may be a sentence containing the phrase to be translated or may be either one of a clause, a phrase, a stereotyped expression, or a collocation containing the phrase to be translated. When either one of a clause, a phrase, a stereotyped expression, or collocation is assumed to be an original sentence to be processed, it is possible to allow the clause, phrase, stereotyped expression, or collocation as the original sentence to be processed to be specified by a user or to be automatically determined through the translation identifying processing.

[0054] In the next Step 102, a natural sentence completely coinciding with the original sentence to be processed is retrieved from natural sentences registered in the corresponding translation database using text data of the original sentence to be processed acquired in Step 100 as a key, and a natural sentence containing the phrase to be translated (a natural sentence partially coinciding with the original sentence because at least the phrase to be translated is contained) is retrieved from the natural sentences registered in the corresponding translation database using text data of the phrase to be translated as a key.

When the phrase to be translated is composed of a plurality of words, natural sentences containing at least one of a plurality of words composing the phrase to be translated is retrieved at the same time.

Step 102 corresponds to the retrieval means described in Claims 1, 2, 15, and 16 and also corresponds to the first step described in Claims 10 and 11.

Step 104 through Step 112 corresponds to the first translation identifying means described in Claims 4.

[0055] In Step 104, whether a natural sentence completely coinciding with the original sentence to be translated is extracted from the corresponding translation database by the retrieval in step 102 or not is asked.

If the answer is positive, the program moves to Step 106 to read the corresponding translations in the target language correlated with a natural sentence completely coinciding with the original sentence to be processed and registered in the corresponding translation database and to display the corresponding translations on display 12 to end the processing.

When displaying these corresponding translations, the read translation is searched for the translation of the phrase to be translated and the translation of the found phrase to be translated is highlighted.

In this case a user can recognize the suitable translation of the specified phrase to be translated (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language) and can recognize the suitable translation of the original sentence to be processed containing the specified phrases to be translated (translation sentences natural as sentences in the target language).

The abovementioned Steps 104 and 106 correspond to the first translation identifying means described in Claim 2.

[0056] (Example 1) The abovementioned processing is further described using an actual example.

Assuming that a source language is English and a target language is Japanese, when “For your safety” contained in an original sentence to be processed “For your safety, don’t rush into the train” is specified as a phrase to be translated, the translation “安全” is selected against “safety” that composes the abovementioned phrase to be translated because the source language and the corresponding translation are registered word by word in a general translation process.

However, in the translation identifying processing relating to this embodiment, since natural sentences in the source language and their corresponding translations in the target language are registered in the corresponding translation database, there is a possibility that natural sentences completely coinciding with the abovementioned original sentences are registered in the corresponding translation database.

When natural sentences completely coinciding are registered in the corresponding translation

database, the translations registered in the corresponding translation database correlated with the

number of coinciding words to be counted.

Thus, the influence of the frequently appearing words on the number of coinciding words can be avoided.

[0060] In this embodiment, when installing the translation-identifying program, an inflected word table where words having different endings due to the difference of singular/plural or tense are registered is also stored in HDD 18.

When the number of coinciding words is counted in Step 108 and the words having only mismatching endings appear, whether the reason for a mismatching ending exists in the difference of the singular/plural or in the difference of the tense is asked, and the words having different endings due to the difference of the singular/plural or the tense are counted assumed as the coinciding words.

Instead of this, the system in which only irregular inflections are registered in the inflection word table and regular inflections (for example, “s” for plural form and “ed” for past tense) are automatically assumed to be the coinciding words is acceptable.

Thus, the influence of the words having different endings due to the singular/plural or the tense (words that should be fundamentally recognized as the coinciding words) on the number of coinciding words can be avoided.

The abovementioned processing in Step 108 corresponds to the first translation identifying means described in Claim 7.

[0061] In this embodiment, a word once counted as the coinciding words for the reason that the word coincides with a certain word in the original sentence during counting of the number of coinciding words in Step 108 are not counted twice even if the word reappears in a natural sentence, in order to avoid repetitively counting the coinciding words that appeared twice or more. Thus even if the same coinciding word exists in multiple places of the natural sentence, the influence of this coinciding word on the number of coinciding words can be avoided.

[0062] A formula for calculating the degree of coincidence can be defined so that the degree of coincidence increases as the number of coinciding words increases. For example, a calculation formula for normalizing the number of coinciding words based on the number of words

composing the phrases to be translated (degree of coincidence = number of coinciding words divided by number of words composing phrases to be translated) can be used.

[0063] In the next Step 110, the degrees of coincidence calculated in each natural formula are compared to ask whether a plurality of natural sentences having the maximum degree of coincidence exists or not.

If the answer is negative, the program moves to Step 112 to read the corresponding translation sentence in the target language registered in the corresponding translation database having a corresponding natural sentence of the maximum degree of coincidence, to search the read translation sentence for the translation of the phrase to be translated, to display the read translation sentence on display 12 with the translation of the found phrase to be translated highlighted, and to end the processing.

In this case, the user can recognize a suitable translation of the specified phrase to be translated (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language).

[0064] In Step 112, in addition to the natural sentences having the maximum degree of coincidence among the natural sentences extracted by the retrieval in Step 102, the translated sentences of a plurality of natural sentences (a certain number of natural sentences in a descending order of the degree of coincidence, or all the natural sentences of which degree of coincidence is higher than the specified value) are also read from the corresponding translation database in a descending order of the degree of coincidence and displayed on display 12 at the same time.

If a plurality of natural sentences having the same degree of coincidence exists, the number of words mismatched with the original sentence in each natural sentence is counted, the translated sentence corresponding to the natural sentence having the same degree of coincidence is displayed in ascending order (from smaller) of the number of mismatched words in the corresponding natural sentences to display the translated sentence at the same time.

[0065] (Example 2) The abovementioned processing is further described using an actual example. Assuming that a source language is English and a target language is Japanese, when a word "operation" contained in an original sentence to be processed, "The system suddenly came into operation," is specified and natural sentences (1) through (8) shown in Figure 3 are correlated with

[0069] When a plurality of natural sentences having the maximum degree of coincidence exists, the answer to the question in Step 110 is positive. Then the program moves to Step 114 to perform a correlation analysis for selecting a natural sentence based on the phrases to be translated and the words that appear frequently in the same sentences of the source language (words having a high correlation with the phrase to be translated: corresponding to the frequently appearing phrases described in Claims 1, 10 and 15) in Steps 114 and 116.

[0070] That is to say, in Step 114, whether the phrase to be translated and the words frequently appearing in the same sentence of the source language (words having a high correlation with the phrase to be translated) are contained in the original sentence to be processed or not is investigated by investigating whether each word other than the phrase to be translated in the original sentence to be processed (however, the frequently appearing words registered in the frequently appearing word table are excluded) is contained in each natural sentence extracted from the corresponding translation database by the retrieval in Step 102 or not.

In this processing, for example, the words having the number of times of appearing in each natural sentence greater than the specified number of times (once, twice or more) among each word other than the phrases to be translated in the original sentences to be processed can be identified as the word having a high correlation to the phrase to be translated.

The abovementioned Step 114 corresponds to the search means described in Claims 1 and 15 and corresponds to the second Step described in Claim 10.

[0071] In the next Step 116, whether the words having a high correlation with the phrase to be translated are found by the retrieval in Step 114 or not is asked.

When the answer is positive, the program moves to Step 118 to refer to the translated sentences of the phrase to be translated and the natural sentences each of which contains the words having a high correlation with the phrase to be translated determined in Step 114 (words existing in the original sentence to be processed) among each natural sentence extracted from the corresponding translation database by the retrieval in Step 102, and to identify the frequently appearing translation of the phrase to be translated in the abovementioned natural sentences.

Specifically speaking for the frequently appearing translation of the phrase to be translated, when the phrase to be translated and a single natural sentence is extracted as a natural sentence containing each of the words having a high correlation with the phrase to be translated, the translation of the phrase to be translated in the natural sentences is identified as the frequently

appearing translation. When a plurality of natural sentences is extracted as the abovementioned

natural sentences, the translation of the most frequently appearing in the translations of the phrases to be translated in these natural sentences are identified as the frequently appearing translations.

[0072] As mentioned above, when natural sentences containing the phrases to be translated and containing words identical to the original sentence to be processed as the words having a high correlation with the phrases to be translated exist, there is a high probability that the natural sentence is a sentence that uses the phrase to be translated as a sentence having the same meaning as the original sentence. There is also a possibility that a natural sentence having different translations of the phrase to be translated is mixed in the abovementioned natural sentence.

On the other hand, in Step 118, the frequently appearing translations of the phrases to be translated in the abovementioned natural sentences are identified by referring to the translated sentences of natural sentences containing each of the phrases to be translated and the words having a high correlation with the phrase to be translated. Therefore, the suitable translations of the phrases to be translated in the original sentences to be processed can be obtained.

[0073] In Step 119, the translated sentences in the target language registered in the corresponding translation database correlated with the natural sentences in which the phrases to be translated are correlated with frequently appearing translations found in Step 118 are read from the natural sentences containing each of the phrases to be translated and the words having a high correlation with the phrases to be translated, and the read translated sentences are displayed on display 12 to end the processing with the translations of the phrases to be translated and the words having a high correlation with the phrases to be translated highlighted on the read translated sentences.

Also, in this case, a user can recognize a suitable translation of the specified phrase to be translated (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language).

The abovementioned Steps 116 through 119 correspond to the first translation identifying means described in Claim 1 and 15, and the third Step described in Claim 10, respectively.

[0074] Also in Step 119, in addition to the natural sentences containing each of the phrases to be translated and the words having a high correlation with the phrases to be translated and where the phrases to be translated are correlated with the frequently appearing translations among the natural sentences extracted by the retrieval in Step 102, the translations of a plurality of natural sentences (a certain number of natural sentences in the descending order of the degree of coincidence, or all the sentences having the degree of coincidence higher than the specified value) are read from the

natural sentences based on the alternative words, i.e., equivalent to the alternative phrases as defined in Claim 2, 3, 11 and 16 that can replace with focus words other than the target phrases to be rendered in the original text to be processed.

[0080] That is to say, the word existing in the original sentence to be processed in Step 120 and not existing in each natural sentence extracted from the corresponding translation database by the retrieval in Step 102 (targeted word) is identified.

In Step 122, a natural sentence containing the targeted word is retrieved from the natural sentences registered in the corresponding translation database.

When a plurality of the targeted words exists, the retrieval in Step 122 is performed for each word. In Step 124, a retrieval condition for retrieving each of the natural sentences (“natural sentence having the same sentence structure as the natural sentence extracted by the retrieval of natural sentences containing the targeted phase” described in Claim 3) only of which targeted word in each natural sentence extracted by the retrieval in Step 122 differs is defined to retrieve the natural sentence from the natural sentences registered in the corresponding translation database based on the defined retrieval condition.

[0081] In the next Step 126, whether the matching natural sentence is extracted by the retrieval in Step 124 or not is asked.

When each of the same translated sentences exists in the corresponding translation database only with the targeted word replaced with the other word, there is a high probability that the aforementioned other word is an alternative word that can be used instead of the targeted word. Therefore, when the answer is positive, the program moves to Step 128 to recognize the word replaced with the targeted word in the natural sentences extracted by the retrieval as the alternative word for the targeted word.

The abovementioned Steps 120 through 130 correspond to the identifying means described in Claim 2 (more specifically in Claim 3) and Claim 16, and the second Step described in Claim 11.

[0082] In Step 130, whether a natural sentence in which the targeted word is replaced with the alternative word exists in a plurality of natural sentences having the maximum degree of coincidence (this may be substituted by “degree of coincidence more than the specified value”) or not is asked.

When a natural sentence in which targeted word is replaced with the alternative word exists in a plurality of natural sentences having the maximum degree of coincidence, it is determined that

there is a high probability that the natural sentence concerned is a sentence that uses the phrase to

be translated in the same meaning as the original sentence to be processed.

Therefore, if the answer is positive, the program moves to Step 132 to read the translated sentence of the target language registered in the corresponding translation database correlated with the natural sentence having the maximum degree of coincidence (this may be substituted by “degree of coincidence more than the specified value”) and containing the alternative word instead of the targeted word, to search the read translated sentence for the translations of the phrase to be translated and of the alternative word, to display the read translated sentence on display 12 with the translation of the recognized phrase to be translated highlighted and with the translation of the alternative word marked so as to be recognized also as an alternative word, and to end the operation.

Also, in this case, a user can recognize a suitable translation of the specified phrase to be translated (the translation having high probability of obtaining from the original sentence a natural translation as a sentence in the target language).

The abovementioned Steps 130 and 132 correspond to the first translation identifying means described in Claim 2 and Claim 16, and the third Step described in Claim 11.

[0083] Also in Step 132, in addition to the natural sentences having the maximum degree of coincidence and containing the alternative word instead of the targeted word among the natural sentences extracted by the retrieval in Step 102, the translated sentences of a plurality of natural sentences (a certain number of natural sentences in the descending order of the degree of coincidence, or all the natural sentences having a degree of coincidence more than the specified value) are read in the descending order of the degree of coincidence and listed on display 12 in the order according to the degree of coincidence and the number of mismatched words at the same time.

[0084] (Example 6) The abovementioned processing is further described using an actual example. Assuming that a source language is English and a target language is Japanese, when a word “have” is specified as the phrase to be translated in the original sentence to be processed containing “have lunch,” a natural sentence containing “have breakfast” (a natural sentence that uses the phrase to be translated “have” in the same meaning as the original sentence to be processed) is registered in the corresponding translation database, a natural sentence containing “have lunch” is not registered, and an appropriate translated sentence cannot be identified even by investigation based on the degree of coincidence or a correlation analysis, the phrase to be

translated in the original sentence to be processed (“have”) or each word other than the words that

translated.

In this case, a user can recognize some candidates for suitable translations of the specified phrase to be translated.

[0091] In the abovementioned examples, although the mode for calculating the degree of coincidence using only the number of coinciding words is described, modes are not limited to this. It is acceptable to define the formula for calculating the degree of coincidence so that the degree of coincidence increases as the number of coinciding words between natural sentences and original sentences increases in order to calculate and evaluate the degree of coincidence based on the number of coinciding words and the number of mismatched words.

It is also acceptable to first evaluate the degree of coincidence only based on the number of coinciding words, to count the number of mismatched words when it is difficult to select a single natural sentence (translated sentence) even by performing a correlation analysis or a scheme analysis, and to select a natural sentence (translated sentence) having the minimum number of words mismatching the original sentence.

[0092] In evaluating the degree of coincidence, it is also acceptable to commonly use a degree of similarity of arranged order of a natural sentence and an original sentence, or the number of mismatched words in natural sentences existing in the words coinciding with the original sentence, in addition to the abovementioned number of coinciding words and number of mismatched words to evaluate the degree of coincidence so that the degree of coincidence with the original sentence increases as the similarity of arranged order of words between a natural sentence and an original sentence increases or to evaluate the degree of coincidence so that the degree of coincidence with the original sentence increases as the number of mismatched words existing between the coinciding words decreases.

The evaluation of the degree of coincidence considering the similarity of arranged order of words between a natural sentence and an original sentence and the number of mismatched words existing between words coinciding with the original sentence can be specifically realized by, for example, the following processing:

Natural sentence 2: $10.0 \text{ (word D)} + 0 \text{ (word X)} + 5.0 \text{ (word E)} + 2.0 \text{ (word F)} + 1.0 \text{ (word G)} = 18.0$

Natural sentence 3: $10.0 \text{ (word D)} + 5.0 \text{ (word E)} + 2.0 \text{ (word F)} + 1.0 \text{ (word G)} = 18.0$

Natural sentence 4: $1.0 \text{ (word A)} + 2.0 \text{ (word B)} + 0 \text{ (word X)} + 5.0 \text{ (word C)} + 0 \text{ (word X)} + 10.0 \text{ (word D)} = 18.0$

Natural sentence 5: $10.0 \text{ (word D)} + 5.0 \text{ (word E)} + 2.0 \text{ (word F)} + 1.0 \text{ (word G)} + 0.5 \text{ (word H)} + 0.2 \text{ (word J)} = 18.7$,

where natural sentence 5 of the sum (the degree of coincidence) of the first assessment value is the largest.

[0095] The first assessment value is defined so that the values increases as the difference between the phrase to be translated in the original sentence and each word other than the phrase to be translated in the original sentence decreases. Therefore by evaluating each natural sentence containing the phrase to be translated based on the sum of the first assessment value as mentioned above, the natural sentence containing many words existing in the position close to the phrase to be translated in the original sentence, in other words, the natural sentence having high probability of containing many phrases composed of the phrase to be translated in the original sentence and the words located close to the phrase to be translated (the natural sentence of which similarity of arranged order of words with the original sentence is estimated to be high) can be evaluated as the natural sentence having higher degree of coincidence.

This mode corresponds to the invention according to Claim 5, 12 and 17.

[0096] Furthermore, the coincidence, i.e., total value of the first evaluated value, in the natural sentences 1-4 is equivalent in the above case, and when there are multiple natural sentences with equivalent coincidence based on the first evaluated value as in this case, the second evaluated value is granted according to the interval, i.e., the number of words, from the target phrase for rendition in the respective natural sentences toward the word corresponding to either of the words in the original text among the respective words in the natural sentence regarding the respective natural sentences with equivalent coincidence based on the first evaluated value.

This second assessment value can also be defined so that the second assessment value for the phrase to be translated contained in each natural sentence is the largest and the value decreases as the difference from the phrase to be translated in each natural sentence increases (as the number of words existing in-between increases).

Thus the sum of the second assessment value provided to each coinciding word in each natural sentence having the same value of the degree of coincidence based on the first assessment value is

calculated, and the translation of each natural sentence having the same value of the degree of coincidence based on the first assessment value is displayed in the descending order of the sum of the second assessment value of each correlated natural sentence (this is also included in the degree of coincidence relating to this invention).

[0097] For example, each coinciding word in the aforementioned natural sentences 1 through 4 is provided with the second assessment value to obtain the sum of the second assessment value as shown below.

In the example below, the second assessment value of the phrase to be translated is assumed to be 10.0, and the second assessment value is defined so that the second assessment value of the other coinciding words is assumed to decrease like 5.0, 2.0, 1.0, 0.5, 0.2, ... as the number of words existing between the phrases to be translated increases like 0, 1, 2, 3, 4,

Natural sentence 1: $10.0 \text{ (word D)} + 0 \text{ (word X)} + 0 \text{ (word X)} + 1.0 \text{ (word E)} + 0.5 \text{ (word F)} + 0.2 \text{ (word G)} = 11.7$

Natural sentence 2: $10.0 \text{ (word D)} + 0 \text{ (word X)} + 2.0 \text{ (word E)} + 1.0 \text{ (word F)} + 0.5 \text{ (word G)} = 13.5$

Natural sentence 3: $10.0 \text{ (word D)} + 5.0 \text{ (word E)} + 2.0 \text{ (word F)} + 1.0 \text{ (word G)} = 18.0$

Natural sentence 4: $0.2 \text{ (word A)} + 0.5 \text{ (word B)} + 0 \text{ (word X)} + 2.0 \text{ (word C)} + 0 \text{ (word X)} + 10.0 \text{ (word D)} = 12.7$

Therefore, natural sentences 1 through 4 (translated sentences) are displayed in the descending order of the sum of the second assessment value, namely in the order of natural sentences 3, 2, 4, 1.

[0098] The second assessment value is defined so that the value increases as the difference between the coinciding word coinciding with any word in the original sentence among the words in each natural sentence and the phrase to be translated in each natural sentence decreases. Therefore the natural sentence containing a larger number of the words coinciding with the original sentence and containing the mismatched words existing between the coinciding words as small as possible, namely the natural sentence having a high probability that the natural sentence contains the phrase to be translated in the original sentence and the phrases composed of the words close to the phrase to be translated. This natural sentence can be evaluated as one with higher coincidence.

Incidentally, the above aspect is also compatible with the inventions as defined in claims 5 (see

claim 6 for details), 12, and 17.

[0099] The evaluation of the similarity of arranging order of words in natural sentences and

8. Natural sentence (h) = (B, C, /, C, D)

9. Natural sentence (j) = (B, X, C, /, C, X, D)

10. Natural sentence (k) = (B, X, X, C, /, C, X, X, D)

By providing each natural sentence containing the phrase to be translated with the third assessment value as shown above and evaluating based on the sum of the third assessment value, the natural sentence having high similarity of arranging order of word with the original sentence and having the number of mismatched words as small as possible existing between the words coinciding with the original sentence, that is to say, the natural sentence containing the original sentence in the condition that the natural sentence contains the mismatched words as small as possible (in the condition of high integrity)) can be evaluated as the natural sentence having a higher degree of coincidence.

This mode corresponds to the invention according to Claim 7, 13 and 18.

[0106] The evaluation considering the similarity of arranged order of words between the natural sentence and the original sentence and the number of mismatched words in the natural sentence existing between the words coinciding with the original sentence can be realized by, for example, the following processing:

[0107] First, the upstream word having the shortest distance from the phrase to be translated (this time the upstream word adjacent to the phrase to be translated) is extracted from the upstream word group existing in the upstream side of the phrase to be translated in the original sentence, then the position in which the phrase to be translated exists in the natural sentence is assumed to be the reference position in each natural sentence containing the phrase to be translated extracted from the corresponding translation database by the retrieval (when the phrase to be translated exists in each of a plurality of places in the natural sentence, the position in which any one of the phrase to be translated exists is assumed to be the reference position), and the region within the specified number of words (for example three words) from the reference position to the upstream side is searched for the upstream word previously extracted.

This processing counts the number of words by distance (more specifically the first number of words by distance and the second number of words by distance) as a fourth assessment value from each of the natural sentences containing the phrase to be translated, and measures the distance between the reference position and the upstream word in the natural sentence in which the upstream word was found in the abovementioned retrieval to determine the first number of words

by distance according to each measured distance.

[0108] The upstream word not yet extracted and having the shortest distance from the phrase to be

8. Natural sentence (h) = (B, X, C, /, C, X, D)

9. Natural sentence (j) = (B, X, X, C, /, C, X, X, D)

In the abovementioned processing, the counting of the second number of words by distance is not essential. When it is not necessary to evaluate the natural sentence for which words are partially rearranged from the original sentence, it is acceptable to omit the counting of the second number of words by distance and use the first number of words by distance as the final evaluation.

Incidentally, the above aspect corresponds to the inventions as defined in claims 8, 14, and 19, and especially, the aspect using the number of words by the first distance in combination with the second accommodates the invention as defined in claim 9.

[0126] It is a matter of course that the evaluation of degree of coincidence and selection or rearrangement of natural sentences based on the abovementioned first assessment value and the second assessment value/third assessment value/fourth assessment value (the first number of words by distance and the second number of words by distance) can be commonly used with evaluation, selection, or rearrangement of natural sentences by a correlation analysis or a scheme analysis. The mode for evaluating the degree of coincidence of natural sentences based on the first assessment value and the second assessment value or the third assessment value is more suitable for the case when retrieving the translated sentence to be used for reference as for the phrases within the range narrower than the entire original sentence, such as phrases existing in the original sentences in comparison with the modes, etc., for evaluating the degree of coincidence using only the number of coinciding words between the original sentence and the natural sentence. For example, it is acceptable to provide a retrieval mode, such as a phrase retrieving mode, in addition to the general retrieval mode, and to perform the evaluation of the degree of coincidence based on the first assessment value and the second assessment value or the third assessment value and the selection or rearrangement of the natural sentences when this phrase retrieval mode is selected.

[0127] Although an example wherein the degree of coincidence is first calculated to select a natural sentence (identify the corresponding translation), a natural sentence is selected (corresponding translation is identified) by a correlation analysis when the suitable natural sentence cannot be narrowed only based on the degree of coincidence, and a natural sentence is selected (corresponding translation is identified) by a scheme analysis when the suitable natural sentence cannot be narrowed down even based on the correlation analysis was described above, the application is not limited to this. It is acceptable to calculate the degree of coincidence for

individual natural sentences containing the phrase to be translated extracted by the retrieval of the retrieval means, to evaluate the individual natural sentences by performing the correlation analysis and the scheme analysis, and to select a natural sentences according to the priority order, for example, shown in Table 6 below based on the evaluation results.

CLAIMS

[1] (After revision)

A system for identifying a corresponding translation, comprising a storage means for storing a plurality of natural sentences, of original sentences composed of a plurality of words, correlated with translated sentences in a target language, a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in a source language from a plurality of natural sentences in said source language stored in said storage means, a search means for searching natural sentences extracted by retrieval of said retrieval means for said phrase to be translated and frequently appearing translations appearing in the same sentence of said source language, a second translation identifying means for searching translated sentences of natural sentences containing each of said phrase to be translated and specific frequently appearing phrases for frequently appearing translations of phrase to be translated by referring to said specific frequently appearing phrases found by said search means and existing in said original sentences and referring to translated sentences of natural sentences containing each of said phrase to be translated among natural sentences extracted by retrieval of said retrieval means and for identifying found frequently appearing translations as translations of said phrases to be translated in said original sentences.

[2] (After revision) A system for identifying a corresponding translation, comprising a storage means for storing a plurality of natural sentences in a source language composed of a plurality of words correlated with translated sentences in a target language, a retrieval means for retrieving natural sentences containing phrases to be translated in original sentences in a source language from a plurality of natural sentences in a source language stored in said storage means, an identifying means for identifying alternative phrases existing in said original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said retrieval means, and a translation identifying means for identifying translations of at least said phrase to be translated in translated sentences of natural sentences containing each of alternative phrases identified by said identifying means

and said phrase to be translated among natural sentences extracted by retrieval of said retrieval means, as translations of at least said phrase to be translated in said original sentences.

[3] (After revision) The system for identifying a corresponding translation according to Claim 2, wherein said identifying means searches a plurality of natural sentences stored in said storage means for natural sentences containing said targeted words, searches natural sentences stored in said storage means for natural sentences having the same sentence structure as natural sentences extracted by said searching, and identifies phrases replaced with targeted phrases in natural sentences extracted by said retrieval as said alternative phrases.

[4] (After revision) The system for identifying a corresponding translation according to Claim 1 or Claim 2, wherein said identifying means determines a degree of coincidence between the natural sentences extracted by the retrieval of the aforementioned retrieval means and the aforementioned original sentence, and identifies the translations of at least the aforementioned phrases to be translated in the translations of the natural sentences selected on the basis of the determined degree of coincidence as the translations of at least the aforementioned phrases to be translated in the aforementioned original sentence.

[5] (After revision)

The parallel translation identifying system consists of the following units: the memorizing means storing multiple natural sentences composed of multiple words in a source language by matching the parallel sentences in a target language; the search unit retrieving a natural sentence containing the object phrases to be rendered in the original text in a source language from among multiple natural sentences in a source language stored in memory; and the parallel translation identifying tool recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the computed total value of the first evaluated value as at least a parallel equivalent for the object phrases to be rendered in the original text, respectively granting the first evaluated value concordant with the small distance in the original text from the object phrases to be rendered in the original

text to the respective words in the original text, recognizing the word corresponding to the original text from among the respective words in the natural sentence in question regarding the natural sentence extracted through a process by the search unit,, thus, computing by the natural sentence the total value of the first evaluated value granted to the identified concordant words.

[6] (After revision) The system for identifying a corresponding translation according to Claim 5, wherein said translation identifying means identifies the parallel translation of the minimal phrases to be rendered in the parallel translation of the natural sentence selected based on the computed total value of the second evaluated value as the parallel translation for the minimal phrases in the original text, if there are multiple sentences from the computed result of the first evaluated value, about the respective natural sentences of such multiplicity, respectively assigning the second evaluated value according to the small distance in the above natural sentence from the object phrases for rendering in the natural sentence for the matching words in the said sentence, computing the total value of the second evaluated value given to the corresponding words aforementioned by the natural sentence,

[7] (After revision) The parallel translation identifying system consists of the following units: the memorizing means storing multiple natural sentences composed of multiple words in source language by matching with the parallel sentences in target language; the search unit retrieving a natural sentence containing the object phrases to be rendered in the original text in source language from among multiple natural sentences in source language stored in the memory; and the parallel translation identifying tool recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the computed total value of the 3rd evaluated value as at least a parallel equivalent for the object phrases to be rendered in the original text, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior

words extracted in the preceding search process within the prescribed number of words from among the anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of granting the 3rd evaluated value concordant with the small interval between the detected specific anterior words when detected and the object phrases to be rendered or the anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the original text, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of granting the 3rd evaluated value concordant with the small interval between the detected specific posterior words when detected and the object phrases to be rendered or the posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the original text, and thus, computing by the natural sentence the total value of the 3rd evaluated value.

[8] (After revision) a storage means for storing a plurality of natural sentences, of original sentences composed of a plurality of words, correlated with translated sentences in a target language

The parallel translation identifying system consists of the following units: the search unit retrieving a natural sentence containing the object phrases to be rendered in the original text in a source language from among multiple natural sentences in a source language stored in the memory; and the parallel translation identifying tool recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the total value computed from the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side as at least the parallel

version for the object phrases in the original text, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the first anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the front side concordant with the interval from the object phrases to be rendered upon detection of the specific anterior words and from the anterior words extracted in the preceding search process till the non-retrieved anterior words have ceased to exist in the first anterior group of words, and concomitantly, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the first posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the back side concordant with the interval from the object phrases to be rendered upon detection of the specific posterior words and from the posterior words extracted in the preceding search process till the non-retrieved posterior words have ceased to exist in the first posterior group of words.

[9] (After revision) The system for identifying a corresponding translation according to Claim 8, wherein said translation identifying means identifies at least the parallel version of the object phrases to be rendered in the parallel sentence of the natural sentence of the result with shorter intra-word distance in the computed integrated recognition as the parallel translation at least of the target phrases to be rendered in the original text, searching, regarding the natural sentence extracted by the search unit, to see whether the specific front words non-retrieved

and with minimal distance from the object phrases to be rendered among the second anterior group of words located on the front compared to the object phrases to be rendered in the original text, excluding the anterior words adjacent to the object phrases to be rendered are aligned on the front of the object phrases to be rendered in the natural sentence or anterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up 1 to the distance between the specific front words when discovered and the object phrases to be rendered or the front words extracted in the preceding search or adding the number of words by the second interval in relation to the corresponding front side to the distance added to the interval computed regarding the front-side words discovered in the preceding search until the non-retrieved front words have ceased to exist in the original text, concomitantly selecting as the final recognition the result of a shorter distance between the respective words from among the number of words by the first and second intervals regarding the front side respectively, and also, searching, regarding the natural sentence extracted by the search unit, to see whether the specific back-side words non-retrieved and with minimal distance from the object phrases to be rendered among the second posterior group of words located on the back side from the object phrases to be rendered in the original text, excluding the posterior words adjacent to the object phrases to be rendered are aligned on the back of the object phrases to be rendered in the natural sentence or posterior group of words extracted in the preceding search and within the prescribed number of words, repeatedly counting up 1 to the distance between the specific back-side words when discovered and the object phrases to be rendered or the back-side words extracted in the preceding search or adding the number of words by the second interval in relation to the corresponding back side to the distance added to the interval computed regarding the back-side words discovered in the preceding search until the non-retrieved back-side words have ceased to exist in the original text, concomitantly selecting as the final recognition the result of a shorter distance between the respective words from among the number of words by the first and second intervals regarding the back side respectively, performing a

comprehensive evaluation by adding up the final recognition on the anterior and posterior sides.

[10] (After revision) The translation identifying method, comprising a first step for retrieving natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language, a second step for searching the natural sentences extracted by the retrieval of the retrieval means in the aforementioned first step for the aforementioned phrases to be translated and the frequently appearing phrases that frequently appear in the same sentences of the source language, and a third step for searching for the aforementioned phrases to be translated and the frequently appearing translations for the phrases to be translated in the corresponding translation sentences of the natural sentences containing each of the aforementioned specific frequently appearing phrases and identifying the found frequently appearing translations as the translations of the aforementioned phrases to be translated in the original sentence by referring to the specific frequently appearing phrases found in the aforementioned second step and existing in the aforementioned original sentence among the natural sentences extracted by the retrieval in the aforementioned first step and referring to the corresponding translation sentences of the natural sentences each of which contains the aforementioned phrases to be translated.

[11] (After revision) The translation identifying method, comprising a first step for retrieving natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language, a second step for identifying alternative phrases existing in said original sentences and replaceable with targeted phrases not contained in natural sentences extracted by retrieval of said first step, and a third step for identifying translations of at least said phrase to be translated in translated sentences of natural sentences containing each of

alternative phrases identified by said second step and said phrase to be translated among natural sentences extracted by retrieval of said first step, as translations of at least said phrase to be translated in said original sentences.

[12] (After revision) a first step for retrieving natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language,

The parallel translation identifying methodology includes the second step to identify the parallel version of the object phrases to be rendered at least in the parallel sentence of the natural sentence selected based on the total value of the first evaluated value computed at least as the parallel version of the object phrases to be rendered in the original text, respectively granting the first evaluated value corresponding to the small distance in the original text from the object phrases to be rendered in the original text toward the respective words in the original text, recognizing the words concordant with the original text from among the respective words in the relevant natural sentence in relation to the natural sentence extracted by the search process in the first step, thus, computing by the natural sentence the total value of the first evaluated value granted to the corresponding words identified.

[13] (After revision) a first step for retrieving natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language,

The parallel translation identifying methodology includes the second step to recognize at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the computed total value of the third evaluated value as at least the parallel version for the object phrases to be rendered in the original text, searching, regarding the natural sentence extracted through the process by the search unit in the first step, to see whether the specific anterior words non-retrieved and with minimal distance from the object

phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of granting the third evaluated value concordant with the small interval between the detected specific anterior words when detected and the object phrases to be rendered or the anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the original text, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of granting the third evaluated value concordant with the small interval between the detected specific posterior words when detected and the object phrases to be rendered or the posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the original text, and thus, computing by the natural sentence the granted total value of the third evaluated value.

[14] (After revision) a first step for retrieving natural sentences containing phrases to be translated in the original sentence in the source language from natural sentences in the source language composed of a plurality of words stored in the storage means by correlating the natural sentences with the translations of the target language,

The parallel translation identifying methodology includes the second step to recognize at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side as at least the parallel version for the object phrases to be rendered in the original text, searching, regarding the

natural sentence extracted by the search process in the first step, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the first anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the front side corresponding to the distance from the object phrases to be rendered when the specific anterior words are detected or the from anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the first anterior group of words, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the first posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the back side corresponding to the distance from the object phrases to be rendered when the specific posterior words are detected or the from posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the first posterior group of words.

[15] A program for allowing a computer connected to a storage means storing a plurality of natural sentences in a source language composed of a plurality of words by correlating with translations in a target language to function as a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in said source language from a plurality of natural sentences in said source language stored in said storage means, a search means for searching natural sentences extracted by retrieval of said retrieval means for said phrase to be translated and frequently appearing phrases that frequently appears in the

same sentence of a source language, and a second translation identifying means for searching translated sentences of natural sentences containing each of said phrase to be translated and said specific frequently appearing phrases for frequently appearing translations of phrase to be translated and for identifying found frequently appearing translations as translations of said phrase to be translated in said original sentences by referring to specific frequently appearing phrases found by said search means and existing in said original sentences and referring to translations of natural sentences each of which contains said phrase to be translated.

[16] (After revision) A program for allowing a computer connected to a storage means storing a plurality of natural sentences in a source language composed of a plurality of words by correlating with translations in a target language to function as a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in said source language from a plurality of natural sentences in said source language stored in said storage means, a search means for searching natural sentences extracted by retrieval of said retrieval means for said phrase to be translated and frequently appearing phrases that frequently appears in the same sentence of a source language, and a second translation identifying means for searching translated sentences of natural sentences containing each of said phrase to be translated and said specific frequently appearing phrases for frequently appearing translations of phrase to be translated and for identifying found frequently appearing translations as translations of said phrase to be translated in said original sentences by referring to specific frequently appearing phrases found by said search means and existing in said original sentences and referring to translations of natural sentences each of which contains said phrase to be translated.

[17] (After revision) A program for allowing a computer connected to a storage means for storing a plurality of natural sentences in the source language composed of a plurality of words having corresponding translations in the target language to function as a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in a source language from a plurality of natural sentences in said source language stored in

said storage means, to function as a retrieval means for retrieving natural sentences containing phrase to be translated in original sentences in a source language from a plurality of natural sentences in said source language stored in said storage means,

[18] (After revision) This is the program to allow the parallel translation identifying methodology to function recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the computed total value of the third evaluated value as at least the parallel version for the object phrases to be rendered in the original text, making the methodology work together with other functions: the memorizing means storing multiple natural sentences composed of multiple words in source language by matching the parallel sentences in a target language and the search unit, by means of the connected computer, retrieving a natural sentence containing the object phrases to be rendered in the original text in a source language from among multiple natural sentences in a source language stored in the memory, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of granting the third evaluated value concordant with the small interval between the detected specific anterior words if detected and the object phrases to be rendered or the anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the original text, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating

the process of granting the third evaluated value concordant with the small interval between the detected specific posterior words if detected and the object phrases to be rendered or the posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the original text.

[19] (After revision) This is the program to allow the parallel translation identifying methodology to function recognizing at least the parallel version of the object phrases to be rendered in the parallel natural sentence selected based on the number of words by the first distance regarding the front side and from the number of words by the first distance regarding the back side as at least the parallel version for the object phrases to be rendered in the original text, making the methodology work together with other functions: the memorizing means storing multiple natural sentences composed of multiple words in a source language by matching the parallel sentences in a target language and the search unit, by means of the connected computer, retrieving a natural sentence containing the object phrases to be rendered in the original text in a source language from among multiple natural sentences in a source language stored in the memory, searching, regarding the natural sentence extracted through the process by the search unit, to see whether the specific anterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the front side of the object phrases to be rendered in the natural sentence or of the anterior words extracted in the preceding search process within the prescribed number of words from among the first anterior group of words located on the front side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the front side corresponding to the distance from the object phrases to be rendered when the specific anterior words are detected or the from anterior words extracted in the preceding retrieval process till the non-retrieved anterior words have ceased to exist in the first anterior group of words, and concomitantly, searching to see whether the specific posterior words non-retrieved and with minimal distance from the object phrases to be rendered are aligned on the back side of the object phrases to be rendered in the natural

sentence or of the posterior words extracted in the preceding search process within the prescribed number of words from among the first posterior group of words located on the back side from the object phrases to be rendered in the original text, repeating the process of counting up the number of words by the first distance regarding the back side corresponding to the distance from the object phrases to be rendered when the specific posterior words are detected or the from posterior words extracted in the preceding retrieval process till the non-retrieved posterior words have ceased to exist in the first posterior group of words.